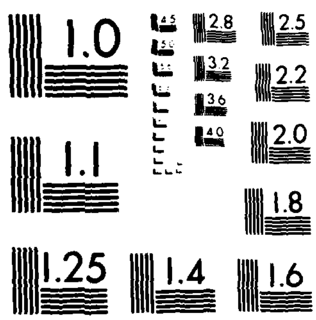
AD-A144 374   ON MAXIMUM ENTROPY SIGNAL ANALYSIS(U) ADMIRALTY MARINE TECHNOLOGY ESTABLISHMENT TEDDINGTON (ENGLAND) L V EMBLING JUN 83 AMTE(N)/TM83057 DRIC-BR-92413   1/1

UNCLASSIFIED   F/G 12/1   NL

END
DATE
FILMED
9 84
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

UNLIMITED

BR92413

(3)

TECH MEMO AMTE(N)TM83057

COPY No 13

AMTE(N)TM83057

ADMIRALTY MARINE TECHNOLOGY ESTABLISHMENT

AD-A144 374

ON MAXIMUM ENTROPY SIGNAL ANALYSIS

L. V. EMBLING

AMTE (Teddington)
Queen's Road TEDDINGTON
Middlesex TW11 0LN

JUNE 1983

UNLIMITED

UX UN~~CLASSIF~~IED

84 08 07 066

UNLIMITED

AMTE(N) TM83057

ON MAXIMUM ENTROPY SIGNAL ANALYSIS

By

L.V. EMBLING

Summary

Maximum entropy signal analysis (MESA) provides a method of deriving the auto power spectrum from a discrete number of time series data; the method is reputed to be superior to Fast Fourier Transform (FFT) analysis when applied to short data records in that better resolution is obtained and sidebands are reduced.

In order to investigate the efficacy of this method MESA has been implemented on computer controlled spectrum analysis systems and a number of simple spectra have been examined. It is concluded that MESA offers no advantages when large data records are available since the quality of the spectra is poor and long computation times are required. Furthermore the method produces highly variable absolute levels. However if only short data records are available then provided a suitable sampling frequency is chosen MESA does provide a reasonable means of producing the power spectrum.

33 pages

17 figures

AMTE (Teddington)
Queens Road
TEDDINGTON, Middx TW11 OLN JUNE 1983

CONTENTS

GLOSSARY OF TERMS

a_{mn}	(n+1)the coefficient of the m+1 long prediction error filter.
E	Expectation value
f	Frequency
FPE	Final Prediction Error
H	Entropy
N	Number of data points
M+1	Order of prediction error filter
P(f)	Power spectrum
P_m	Power output by (m+1)th order prediction error filter
R_i	Measured values of auto-correlation function
R(t)	True values of auto-correlation function
S(f)	Linear spectrum
t	Time
T	Time period of data sample
Δt	Time period between data points
W(t)	Weighting function
$x(t);(x_i)$	Time series data (digitised)

1. INTRODUCTION

Spectral data is generally analysed by means of the auto-power spectrum P(f) of the signal. This is conventionally determined from the measured time signal x(t) by means of the Fourier Transform:

$$
\begin{array}{ccc}
\text{TIME SPECTRUM} & \longrightarrow & \text{AUTO CORRELATION FUNCTION} \\
x(t) & & R(t) = \lim\limits_{T\to\infty} \frac{1}{T} \int x(\tau)x(\tau+t)\, d\tau \\
\Big\downarrow \ \text{i) FILTER } W(t) \ \text{ii) FOURIER TRANSFORM} & & \Big\downarrow \ \text{i) FILTER } W(t) \ \text{ii) FOURIER TRANSFORM} \\
\text{LINEAR SPECTRUM} & \longrightarrow & \text{AUTO POWER SPECTRUM} \\
S(f) = F[x(t)] & & P(f) = F[R(t)] \\
= F[x(t)W(t)] \text{ in practice.} & & = F[R(t)W(t)] \text{ in practice} \\
 & & = S(f)S^{*}(f)
\end{array}
$$

In practice the time series data are measured within some finite time interval and the analysis assumes data outside this interval to be zero. To reduce the error introduced by this assumption the available data is generally treated by some weighting function W(t) such as the Hanning function.

Maximum entropy spectral analysis (MESA) provides an alternative method for determining the power spectrum P(f) from the available data without such modification. In principle the method provides a means of extending the sampled data x_i ($i \leq N$) beyond the measurement window (i.e. to $i > N$) by taking a weighted sum of the known data. The weighting function is known as the Prediction Error Filter and (as shall be explained fully later) is determined by minimising the error power output from a least mean square fit of the measured and 'predicted' data.

The method was first developed by Burg [1] and involves the maximisation of the 'entropy' H of the system with respect to the power spectrum subject to constraints imposed by the measured data. The method is reported to have had much success in improving the power spectrum in that i) better resolution is obtained ii) sidebands are reduced and iii) more realistic power estimates are produced, particularly for short data records (see for example [1] and [2]).

In order to investigate the claims and assess the potential of the method for use on machinery noise data MESA has been implemented on two analysis systems (the GENRAD 2508 Signal analysis system and the Hewlett Packard 5423A Spectral dynamics analyser used un conjunction with the HP 9845 desk top computer) and has been used to analyse several simple spectra.

2. THEORY

The concept of entropy as used in MESA was originally derived from the field of information theory. The entropy of a system is the average rate of information conveyed and is given by

$$H = \sum_{i=1}^{n} p_i \ln(p_i) \qquad 1$$

where p_i is the probability that the ith event of n possible events occurs.

From equation 1 it can be seen that the entropy is minimised when the probabilities p_i are equal, i.e. when the system is most ordered. The application of the concept of entropy to spectral analysis therefore involves the minimisation of the entropy for a stationary Gaussian process which has been shown [1,2,3] to be proportional to

$$H = \int \ln P(f) \; df \qquad 2$$

where P(f) is the power spectrum and must be consistent with the measured values of the auto-correlation function (assuming these to be known exactly).

The minimisation procedure is described in Appendix 1 and results in the following expression for the power spectrum:

$$P(f) = \frac{P_m \, \Delta t}{\left| 1 - \sum_{n=1}^{m} a_{mn} \exp(-2\pi i f n \Delta t) \right|^2} \qquad 3$$

where Δt is the time interval between sampled data points, P_m is the power output by the prediction error filter of order m+1 having elements a_{mn}; these elements are the solutions to the equation

$$\begin{bmatrix} R_0 & R_1 & \cdots & R_m \\ R_1 & R_0 & \cdots & R_{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ R_m & R_{m-1} & \cdots & R_0 \end{bmatrix} \begin{bmatrix} 1 \\ -a_1 \\ \vdots \\ -a_m \end{bmatrix} = \begin{bmatrix} P_m \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad 4$$

where R_k are the measured values of the auto-correlation function.

This analysis is performed on the assumption that the values R_k ($= r_{i-j} = r_{ij} = E[x_i \, x_j]$) of the auto correlation function are exactly known. However, in practice R_k are determined from truncated sampled time series data x_i whence R_k are no longer exact but are given by

$$R_k = \frac{1}{N} \sum_{i=1}^{N-k} x_i \, x_{i+k} \qquad 5$$

Now, by considering the final row of 4 it can be seen that R_k can be expressed in terms of the prediction error filter coefficients and preceeding values of the auto-correlation coefficients:

$$R_k = \sum_{i=1}^{k} R_{k-i}\, a_i \qquad 6$$

Burg [4] proposed a means of obtaining more realistic values of R_k by iterating values for the filter coefficients by successive incrementation of the dimension of the filter. (It is convenient at this stage to redefine the elements a_j as a_{mj} where m refers to the number of iterations performed.) This method is based essentially on a least squares fitting of the true data to the output of the prediction error filter. Consider, for example, the two point prediction error filter

$$\begin{bmatrix} 1 \\ a_{11} \end{bmatrix}$$

In effect, this filter serves to 'predict' the (i+1)th value of the data x_{i+1}(pred) from the ith measured value x_i by setting $x_{i+1}(\text{pred}) = a_{11}x_i$. In order to obtain the 'best' value of a_{11} the filter is scanned along the data sample and a least squares fit is performed between the measured and predicted values to determine the value of a_{11} which minimises the 'error' P_1. In order to ensure that the resulting filter is stable, it is required that the coefficients a_{ij} should be less than unity. To satisfy this criterion Burg [4] suggested that the forward power output of the filter should be averaged with the power output by the filter operated in reverse:
reverse.

$|a_{11} \quad 1|$ Forward filter → $|a_{11} \quad 1|$

$x_{i+1}\text{pred} = a_{11}x_i$

$x_1 \quad x_2 \quad x_3 \quad x_4 \quad \ldots \quad x_{N-2} \quad x_{N-1} \quad x_N$

$|1 \quad a_{11}|$ ← Backward filter $|1 \quad a_{11}|$

$x_i\text{pred} = a_{11}x_{i+1}$

Thus

$$P_1 = \frac{1}{2(N-1)} \sum_{i=1}^{N-1} \left\{ (x_{i+1} - a_{11}x_i)^2 + (x_i - a_{11}x_{i+1})^2 \right\} \qquad 7$$

Minimising P_1 with respect to a_{11}: $\dfrac{dP_1}{da_{11}} = 0$

whence
$$a_{11} = \frac{2 \sum_{i=1}^{N-1} x_i\, x_{i+1}}{\sum_{i=1}^{N-1} (x_i^2 + x_{i+1}^2)} \qquad 8$$

The prediction error filter could become more accurate on increasing the number of data points incorporated in the estimate resulting in an increased filter length. The procedure for calculating a new set of filter coefficients on incrementing the filter length is described in Appendix 2. In general we find the coefficients through the recursion relation

$$a_{mk} = a_{m-1\,k} - a_{mm}\, a_{m-1\,m-k} \qquad k<m \qquad 9$$

with a_{mm} being derived from preceeding coefficients and the original data

A recursion for the error power output can also be derived (Appendix 3) and is given by

$$P_m = P_{m-1} (1 - a_{mm}^2) \qquad 10$$

where, from 4 and 5

$$P_0 = R_0 = \frac{1}{N} \sum_{i=1}^{N} x_i^2 \qquad 11$$

Returning to the expression for the power spectrum, equation 3, it can be seen that knowing the elements a_{kj} and the error power output P_k the explicit derivation of the auto-correlation coefficients becomes unnecessary and solution of 4 and 6 can be by-passed.

3 ORDER OF PREDICTION ERROR FILTER

The choice of length for the prediction error filter still appears to be a matter of concern. Too short a filter results in smoothing of the data while too long a filter causes unwanted splitting of the peaks. Three criteria are available for choosing the length of the filter [6,7]:

i) Final Prediction Error (FPE) Criterion

This procedure was proposed by Akaike [8,9] and consists of determining the filter length M for which the FPE = $E\{(x_t-\hat{x}_t)^2\}$ is minimised. It is shown that

$$FPE(m) = \frac{(N+m+1)}{(N-m-1)} P_m \qquad 12$$

should be minimised.

ii) Information Theoretic Criterion (AIK)

This is based on the minimisation of the log liklihood of the prediction error variance

$$AIC(m) = \ln P_m + \frac{2m}{N} \qquad 13$$

iii) Autoregressive Transfer Function Criterion (CAT)

This is based on minimising the difference between the mean square errors of the estimated and true filters:-

$$CAT(M) = \frac{1}{N} \sum_{m=1}^{N} \frac{N-m}{N P_m} - \frac{N-M}{N P_m} \qquad 14$$

The first of these methods, viz the FPE criterion, has been most widely used in the literature and was adopted in this assessment.

Ulrych and Bishop [2] recommend M=N/2 as a generally suitable filter length; however, from the point of view of computer time a shorter filter is preferable.

4 COMPUTER PROGRAMS

MESA has been implemented on both the GENRAD 2508 Signal analysis system (Time Series Language, TSL) and on the HP9845 desk top computer (BASIC) used in conjunction with the HP 5423A Spectral Dynamics Analyser. Figure 1 shows the overall structure of both programs while Figure 2 shows the algorithm used for determining the power output P_m and the filter coefficients a_{mn}. In both cases the FPE is plotted as a function of m to allow the user to assess the length of the filter used; this length can be changed if required before going on to determine the power spectrum.

On running the program on the GENRAD the levels of the power spectra exhibited high variability; this was initially attributed to rounding errors incorporated in the TSL software. Since the method is particularly sensitive to such errors the majority of plots presented in this study were derived using the Hewlett Packard system where the increased precision available would minimise their effect. However (as described in the next section) the amplitudes were still found to vary considerably.

5. ANALYSIS PERFORMANCE

The MESA method was tested on various signals including

i) Pure sine wave (1v rms, 50Hz)
ii) Two adjacent sine waves (1v rms, 50Hz; 1v rms, 49Hz)
iii) Square wave (1v rms, 10Hz)
iv) Typical machinery noise .

i) 1v 50Hz sine wave

This simple case was explored fairly extensively with the sampling frequency, number of data points and number of iterations being varied.

The general observations from this investigation were as follows:

a) A two point error filter (M=1) was incapable of producing a suitable spectrum.

b) One iteration ($M=2$) was sufficient to produce a peak in the spectrum; increasing the number of iterations did not affect the position of this peak but did improve the resolution and dynamic range. This conclusion is illustrated in Figure 3 which shows spectra derived from 2,8 and 32 iterations on 64 data points sampled at 256Hz. Note that for $M=32(=N/2)$ peak splitting occurred; such splitting for long filters has been reported in the literature [10]. The corresponding variation of FPE with M is shown in Figure 4: there is no obvious minimum but a levelling off occurs at around $M=8$ and this does correspond to a reasonable spectrum.

c) The position and width of the peak varied with the number of data points and with the sampling frequency. This is illustrated in Figure 5 which shows the peak position converging on 50Hz as the number of points in the sample increased (sampling frequency 3200Hz, $M=2$). on decreasing the sampling frequency the convergence became more rapid (Figure 6) such that a true spectrum could be produced with fewer points. Indeed, at the Nyquist frequency an accurate (excepting peak height) spectrum could be produced with as few as three data points (Figure 7); However the reflection of the zero frequency peak at 100Hz shows that aliasing was occurring.

The resolution also improved as the peak converged to the correct frequency.

d) The method was incapable of producing the correct absolute level of the spectrum.

The major conclusions from this simple investigation were that MESA cannot be used to give accurate levels. Nethertheless it is capable of producing a well resolved spectrum (of a sine wave) using very few data points providing the sampling frequency is low; however on approaching the Nyquist frequency aliasing is seen to occur.

ii) Two adjacent sine waves (1v 49Hz & 50Hz)

Figure 8 shows the ME spectra derived from a 128 point data set sampled at 200Hz to avoid aliasing. It can be seen that four iterations were insufficient to resolve the peaks while 8 and 32 iterations resolved two (unequal) peaks at around 49Hz and 51Hz, the latter frequency being incorrect. On decreasing the length of the data sample to 16 points MESA was still capable of resolving two peaks (albeit at 46Hz and 51Hz) when 8 iterations were used; however an 8 point data sample was insufficient for these to be resolved.

The FPE criterion produced similar results as for the single sine wave and did not provide any guidance as to the best filter length.

iii) 10Hz square wave

In order to assess MESA's ability to cope with harmonics, a 10Hz square wave was examined; the majority of tests were performed with a sampling frequency of 200Hz in order to examine harmonics up to 100Hz.

Figure 9 shows the analysis results for a 128 point data sample using different filter lengths. Eight iterations were sufficient to produce a fair representation of the spectrum in that the first four of the odd harmonics occurred close to the correct frequencies although the peak heights were subject to variability. The fifth odd harmonic was only resolved in the case of the long filter. Increasing the number of iterations improved the resolution. For m=32 peaks were observed at even harmonics; since these were also observed on performing a FFT analysis they were attributed to to the imperfection of the input sine wave.

Application of the FPE criterion proved unhelpful in choosing the 'best' filter length to be used; after 32 iterations no true minimum was attained (Figure 10). (Longer filters were not considered due to the impracticality of computation time - more than ten minutes were required to produce a spectrum using N=128, M=32) Figure 11 shows the ME spectra of shorter length samples (M=8 in each case) and shows that even with only 16 data points a representative spectrum can be produced. However Figure 12 indicates the importance of choosing a suitable sampling frequency: Too low a sampling frequency caused aliasing (f_s=100Hz) while at too high a frequency the small data sample only represented a fraction (<1/2) of a cycle.

The conclusions drawn from this test are that

i) MESA is capable of producing a representative spectrum (in terms of peak position) of the first few harmonics of a 10Hz square wave even when using a small data sample. However higher harmonics were not resolved.

ii) The performance is critically sensitive to the sampling frequency.

iii) The resolution can be improved by using larger data samples but computation time becomes impracticable.

iv) Peak levels are incorrect even in relative terms.

v) The FPE criterion offers little guidance in choosing filter length.

iv) Machinery Noise

A 30kW motor was chosen to represent a simple example of machinery noise; the 'true' spectrum (FFT using 512 data points) for a random run is shown in Figure 13 (not corrected for measurement parameters).

The MESA performance on the first 128 of these data points is shown in Figure 14 for 64 (N/2 as recommended in [2]), 20 (where FPE levelled off), and 5 iterations of the prediction error filter. It can be seen that with only 5 iterations MESA was able to identify the two major peaks while with M=20 and M=64 minor peaks were also resolved. However the performance above about 70 Hz was disappointing, a feature also observed on examination of the square wave. Figure 16 illustrates the performance of MESA using only 16 data points and shows that the major peaks could be accurately resolved although minor peaks had dissappeared due to lack of information. A further test using only 8 data points succeeded in resolving the major (30Hz) peak accurately and exhibited a second peak at around 70Hz (instead of 60Hz)

6 COMPARISON WITH FFT FOR FEW DATA POINTS

A short program was implemented on the GENRAD 2508 to determine the auto-power spectrum by applying the FFT to small samples of data points. In order to preserve the resolution of the output spectrum zero packing was utilised (viz the FFT was performed on a 128 point array incorporating N data points, all other elements being set to zero).

Results for the 50Hz sine wave and 10Hz square wave are shown in Figure 17 for 16, 32 and 64 data points, both waves being sampled at 200Hz.

Comparing these figures with those derived using MESA (Figures 5, 7, and 9) it can be seen that the MESA performance is by far the more satisfactory when only 16 or 32 data points are present in that peaks can be easily resolved; for 64 points MESA improves on the dynamic range and resolution.

7 SIGNAL VARIANCE - DISCUSSION

Referring to equation 3 the ME power spectrum can be seen to be the result of

$$P(f) = \frac{\text{Error power spectrum level}}{|\ \text{F.T. of prediction error filter}\ |^2} \qquad 15$$

Thus in order for peaks to be observed in P(f) the Fourier transform must exhibit minima. The peak amplitude is therefore a result of a limiting process in which both the error power and Fourier transform tend to zero but in such a way that the power spectrum level remains finite.

This probably explains in part the large variation observed in the amplitudes. This is also commented on in [2] as being associated with harmonic signals.

A second source of variance arises from rounding errors. These are particularly troublesome in the determination of the error power P_m (equation 10) when a_{mm} tends to unity. This would occur (see equation A2.6) when $b'_{mt} = b_{mt}$ which is effectively saying that forward and backward predictors give the same output as would occur if the sampling points were symmetric in the signal. (This would only be true of a noiseless signal which would be unlikely in practice.)

8 CONCLUSIONS

MESA has been shown to be a viable means of analysing data provided suitable sampling frequencies and filter lengths are chosen. However a major shortcoming is that the levels derived are extremely variable. Frequency shifting can occur but can be reduced by choosing a sampling frequency close to the Nyquist frequency. Spectral resolution is improved by increasing the length of the prediction error filter; however if this filter is too long spurious peaks can be produced.

Akaike's FPE criterion did not prove particularly helpful as a guide to filter length in any of the examples considered - no clear minima occured while the value of M following a rapid decrease in FPE seemed as good a choice as any.

When long data samples are available MESA offers no advantages over the conventional FFT means of analysis: the spectra are of poor quality, do not provide reliable absolute levels, and their execution is unacceptably slow. However, when only few data points are available MESA has definate advantages in that major peaks can be clearly resolved although reservation is recommended with regard their position.

In the context of time series analysis of machinery noise MESA is unlikely to be of practicable value since data samples can generally be made sufficiently long to allow conventional analysis. However for problems involving spatial analysis where long data samples are inconvenient if not impossible, MESA provides a potentially useful tool.

L.V. EMBLING. (HSO)

APPENDIX 1
MAXIMUM ENTROPY POWER SPECTRUM

The entropy for a stationary Gaussian process is proportional to

$$H = \int_{-\infty}^{\infty} \ln[p(f)]\, df \qquad 2$$

where the power spectrum $P(f)$ must be consistent with the measured values R_k ($k=0$ to m) of the auto-correlation function

$$R_k = \int df\, P(f) \exp(2\pi i f k \Delta t) \qquad \text{A1.1}$$

with $R_k = R_{i-j} = R_{ij} = E[x_i x_j]$; $j = i+k$

The procedure for minimising H with respect to P(f) in order to determine P(f) is described fully in [3]. The method consists in introducing Lagrange multipliers Θ_k where $\Theta_k = \Theta^*_{-k}$ and demanding that

$$\delta(L+H) = 0 \qquad \text{A1.2}$$

where
$$L = \sum_{k=-m}^{m} \Theta_k^* R_k \qquad \text{A1.3}$$

Combining A1.1,A1.2 and A1.3,

$$0 = \delta(L+H) = \int df\, \delta(P(f)) \left[\frac{1}{P(f)} - \sum_{k=-m}^{m} \Theta_k^* \exp(2\pi i f k \Delta t) \right]$$

and since P(f) is arbitrary:-

$$P(f) = \left[\sum_{k=-m}^{m} \Theta_k^* \exp(2\pi i f k \Delta t) \right]^{-1} \qquad \text{A1.4}$$

Imposing the condition that P(f) is positive and integrable it can be expressed as

$$P(f) = \frac{1}{A(f)\, A^*(f)} \qquad \text{A1.5}$$

where
$$A(f) = \sum_{k=0}^{m} \gamma_k^* \exp(2\pi i f k \Delta t) \qquad \text{A1.6}$$

and
$$\Theta_n^* = \sum_{k=0}^{m-n} \gamma_k \gamma_{k+n}^* \qquad \text{A1.7}$$

Multiplying both sides of A1.5 by $A^* \exp(2\pi i f k \Delta t)$ and integrating in the upper half plane it is eventually shown by considering the constraint equations A1.1 and using the periodicity of A(f) that

$$\sum_{l=0}^{m} R_{k,l} \gamma_l = \frac{1}{\gamma_0} \delta_{k,0}$$

PREVIOUS PAGE IS BLANK

or, in matrix notation

$$[R][\gamma] = 1/\gamma_0[\delta] \qquad \text{A1.9}$$

where

$$[R] = \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1m} \\ R_{21} & R_{22} & \cdots & R_{2m} \\ \vdots & \vdots & & \vdots \\ R_{m1} & R_{m2} & \cdots & R_{mm} \end{bmatrix} = \begin{bmatrix} R_0 & R_1 & \cdots & R_{m-1} \\ R_1 & R_0 & \cdots & R_{m-2} \\ \vdots & \vdots & & \vdots \\ R_{m-1} & R_{m-2} & \cdots & R_0 \end{bmatrix}$$

$$[\gamma] = \begin{bmatrix} \gamma_0 \\ \gamma_1 \\ \vdots \\ \gamma_{m-1} \end{bmatrix} \quad \text{and} \quad [\delta] = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{A1.10}$$

whence $$[\gamma] = \frac{1}{\gamma_0} [R]^{-1}[\delta] \qquad \text{A1.11}$$

Finally, from A1.5 and A1.6

$$\begin{aligned} P(f) &= |A(f)|^{-2} = |A^*(f)|^{-2} \\ &= \left| \sum_{k=0}^{m-1} \gamma_k \exp(-2\pi i f k \Delta t) \right|^{-2} \\ &= \left| [\gamma]^T [\epsilon(f)] \right|^{-2} \\ &= \gamma_0^2 \left| ([R]^{-1}[\delta])^T [\epsilon(f)] \right|^{-2} \end{aligned} \qquad \text{A1.12}$$

where

$$[\epsilon(f)] = \begin{bmatrix} 1 \\ \vdots \\ \exp(-2\pi i(n-1)f\Delta t) \end{bmatrix} \qquad \text{A1.13}$$

From A1.11 we find

$$\gamma_0 = \frac{1}{\gamma_0}[R]^{-1}_{11} \qquad \text{A1.14}$$

whence

$$P(f) = \frac{P_m \Delta t}{\left|([R]^{-1}[\delta])^T [\epsilon(f)] \right|^2} \qquad \text{A1.15}$$

Rewriting A1.15 in the form used by [1,5 and 6]

$$P(f) = \frac{P_m \Delta t}{\left| 1 - \sum_{n=1}^{m} a_n \exp(-2\pi i f n \Delta t) \right|^2} \qquad 3$$

16

where $P_m \Delta t = \gamma_0^2$ and $a_n = -\gamma_n/\gamma_0$ are solutions of 4 (viz A1.9); P_m is known as the power output of the prediction error filter, $[a_n]$ and $[R]$ is the TOEPLITZ matrix [1]:

$$\begin{bmatrix} R_0 & R_1 & \cdots & R_m \\ R_1 & R_0 & \cdots & R_{m-1} \\ \vdots & \vdots & \ddots & \vdots \\ R_m & R_{m-1} & \cdots & R_0 \end{bmatrix} \begin{bmatrix} 1 \\ a_1 \\ \vdots \\ a_m \end{bmatrix} = \begin{bmatrix} P_m \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad 4$$

$P(f)$ is the MAXIMUM ENTROPY solution to the power spectrum.

APPENDIX 2
ITERATIVE PROCEDURE

The average error power output of a m+1 long prediction error filter $[a_m]^T = [1 \quad a_{m1} \quad \cdots \quad a_{mm}]$ is given by [5]:

$$P_m = \frac{1}{2(N-m)} \sum_{t=1}^{N-m} \left[\left\{ x_t - \sum_{k=1}^{m} a_{mk} x_{t+k} \right\}^2 + \left\{ x_{t+m} - \sum_{k=1}^{m} a_{mk} x_{t+m-k} \right\}^2 \right] \qquad \text{A2.1}$$

The filter coefficients are given by the recurrance relations (see Appendix 3)

$$a_{mk} = a_{m-1\,k} - a_{mm}\, a_{m-1\,m-k} \qquad k<m \qquad 9$$

where $a_{mk} = 0$ for $k>m$.

In order to determine the coefficient a_{mm} we insert 10 in A2.1 to give

$$P_m = \frac{1}{2(N-m)} \sum_{t=1}^{N-m} \left\{ (b_{mt} - a_{mm} b'_{mt})^2 + (b'_{mt} - a_{mm} b_{mt})^2 \right\} \qquad \text{A2.2}$$

where

$$b_{mt} = \sum_{k=0}^{m} a_{m-1\,k}\, x_{t+k} = \sum_{k=0}^{m} a_{m-1\,m-k}\, x_{t+m-k}$$

$$b'_{mt} = \sum_{k=0}^{m} a_{m-1\,k}\, x_{t+m-k} = \sum_{k=0}^{m} a_{m-1\,m-k}\, x_{t+k} \qquad \text{A2.3}$$

Note $\quad b_{1t} = x_t$ and $b'_{1t} = x_{t+1}$ $\qquad$ A2.4

Since b_{mt} and b'_{mt} are independent of a_{mm} the minimisation of P_m with respect to a_{mm} gives

$$\frac{1}{2(N-m)} \sum_{t=1}^{N-m} -4b_{mt} b'_{mt} + 2{b'_{mt}}^2 a_{mm} + 2b_{mt}^2 a_{mm} = 0$$

and hence

$$a_{mm} = 2 \sum_{t=1}^{N-m} b_{mt} b'_{mt} \qquad \sum_{t=1}^{N-m} ({b'_{mt}}^2 + b_{mt}^2) \qquad \text{A2.5}$$

[NB this reduces to equation 9 as required on substituting m=1]

Furthermore, since

$$\frac{d^2P_m}{da_{mm}^2} = \frac{1}{(N-m)} \sum_{t=1}^{N-m} (b'^{\,2}_{mt} + b^2_{mt}) > 0 \qquad \text{A2.6}$$

the extremum is a minimum as required.

Combining A2.1 and A2.2 gives recursion relations for b_{mt} and b'_{mt} :

$$\begin{aligned} b_{mt} &= b_{m-1\ t} - a_{m-1\ m-1} b'_{m-1\ t} \\ b'_{mt} &= b'_{m-1\ t+1} - a_{m-1\ m-1} b_{m-1\ t+1} \end{aligned} \qquad \text{A2.7}$$

APPENDIX 3
RECURSION RELATIONSHIPS

Consider the equation defining the power output P_{m-1} using an m point prediction error filter (cf eq 4):

$$\begin{bmatrix} R_0 & R_1 & \cdots & R_{m-1} \\ R_1 & R_0 & \cdots & R_{m-2} \\ \vdots & \vdots & & \vdots \\ R_{m-1} & R_{m-2} & \cdots & R_0 \end{bmatrix} \cdot \begin{bmatrix} 1 \\ -a_{m-1\ 1} \\ \vdots \\ -a_{m-1\ m-1} \end{bmatrix} = \begin{bmatrix} P_{m-1} \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad \text{A3.1}$$

Increasing the order of the matrices by 1:

$$\begin{bmatrix} R_0 & R_1 & \cdots & R_{m-1} & R_m \\ R_1 & R_0 & & \vdots & \vdots \\ \vdots & \vdots & & \vdots & \vdots \\ R_{m-1} & R_m & \cdots & R_0 & R_1 \\ R_m & R_{m-1} & \cdots & R_1 & R_0 \end{bmatrix} \begin{bmatrix} 1 \\ -a_{m-1\ 1} \\ \vdots \\ -a_{m-1\ m-1} \\ 0 \end{bmatrix} = \begin{bmatrix} P_{m-1} \\ 0 \\ \vdots \\ 0 \\ \Delta_{m-1} \end{bmatrix} \qquad \text{A3.2}$$

Where

$$\Delta_{m-1} = R_m - \sum_{j=1}^{m-1} R_{m-j} a_{m-1\ j}$$

Making use of the symmetry of A3.2 the prediction error filter can be reversed to give:

$$\begin{bmatrix} R_0 & R_1 & \cdots & R_{m-1} & R_m \\ R_1 & R_0 & & & \vdots \\ \vdots & & & & \vdots \\ R_{m-1} & & & & \vdots \\ R_m & \cdots & \cdots & \cdots & R_0 \end{bmatrix} \begin{bmatrix} 0 \\ -a_{m-1\ m-1} \\ \vdots \\ -a_{m-1\ 1} \\ 1 \end{bmatrix} = \begin{bmatrix} \Delta_{m-1} \\ 0 \\ \vdots \\ 0 \\ P_{m-1} \end{bmatrix} \qquad \text{A3.3}$$

The solution for the m+1th power estimate P_m is given by equation 4

$$\begin{bmatrix} R_0 & R_1 \cdots\cdots & R_m \\ R_1 & R_0 & \vdots \\ \vdots & & \vdots \\ R_m & \cdots\cdots & R_0 \end{bmatrix} \begin{bmatrix} 1 \\ -a_{m1} \\ \vdots \\ -a_{mm} \end{bmatrix} = \begin{bmatrix} P_m \\ 0 \\ \vdots \\ 0 \end{bmatrix} \qquad 4$$

Let the prediction error filter be defined in terms of the (m-1)th filter:

$$\begin{bmatrix} 1 \\ -a_{m1} \\ \vdots \\ -a_{mm-1} \\ -a_{mm} \end{bmatrix} = \begin{bmatrix} 1 \\ -a_{m-1\ 1} \\ \vdots \\ -a_{m-1\ m-1} \\ 0 \end{bmatrix} + \Gamma \begin{bmatrix} 0 \\ -a_{m-1\ m-1} \\ \vdots \\ -a_{m-1\ 1} \\ 1 \end{bmatrix} \qquad A3.4$$

Then equating elements on both sides of A3.4 we find

$$-a_{mm} = \Gamma$$

and $$a_{mk} = a_{m-1\ k} + \Gamma\ a_{m-1\ m-k} \qquad i \neq m$$

$$= a_{m-1\ k} - a_{mm}\ a_{m-1\ m-k} \qquad 9$$

The corresponding power output is given by

$$\begin{bmatrix} P_m \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} P_{m-1} \\ 0 \\ \vdots \\ 0 \\ \Delta_{m-1} \end{bmatrix} + \Gamma \begin{bmatrix} \Delta_{m-1} \\ 0 \\ \vdots \\ 0 \\ P_{m-1} \end{bmatrix} \qquad A3.5$$

whence $$P_m = P_{m-1} + \Gamma\ \Delta_{m-1}$$

and $$\Delta_{m-1} + \Gamma\ P_{m-1} = 0$$

Hence $$P_m = P_{m-1} - \Gamma^2\ P_{m-1}$$

$$P_m = P_{m-1}(1-a_{mm}^2) \qquad A3.7$$

REFERENCES

* 1. BURG, J.P.
Maximum Entropy Spectral Analysis
Proc 37th meeting of Society of Exploration Geophysicists, (1967)

* 2. ULRYCH, T.J. & BISHOP, T.N.
Maximum Entropy Spectral Analysis and Autoregressive Decomposition.
Rev. Geophys & Space Physics 13, 183-200, (1975)

* 3. EDWARD, J.A. & FITELSON, M.M.
Notes on Maximum Entropy Processing.
IEEE Transactions on Information Theory IT-19, 232-234 (1973)

* 4 BURG, J.P.
A New Analysis Technique for Time Series Data.
Presented at NATO Advanced Study Institute on Signal Processing with Emphasis on Underwater acoustics. (Aug. 12-23, 1968)

* 5 ANDERSON, N.
On the Calculation of Filter Coefficients for Maximum Entropy Spectral Analysis.
Geophys 39 69-72 (1974)

* 6 LANDERS, T.E. & LACOSS R.T.
Some Geophysical Applications of Autoregressive Spectral Estimates.
IEEE Trans. Geosci. Electron. GE15, 26-32,(1977)

7 HAYKIN, S.
Non-linear Methods of Spectral Analysis.
Topics in Advanced Physics 34
Springer-Verlag Berlin (1979)

8 AKAIKE, H.
Power Spectrum Estimation Through Auto Regressive Model Fitting.
Ann Inst Stat Math 21,407-419,(1969)

9 AKAIKE,H.
Stastistical Predictor Identification.
Ann Inst Stat Math 22, 203-217, (1970)

10 CHILDERS, D.G.
Modern Spectrum Analysis.
IEEE Press 1978

* Reprinted in Reference 10.

PREVIOUS PAGE IS BLANK

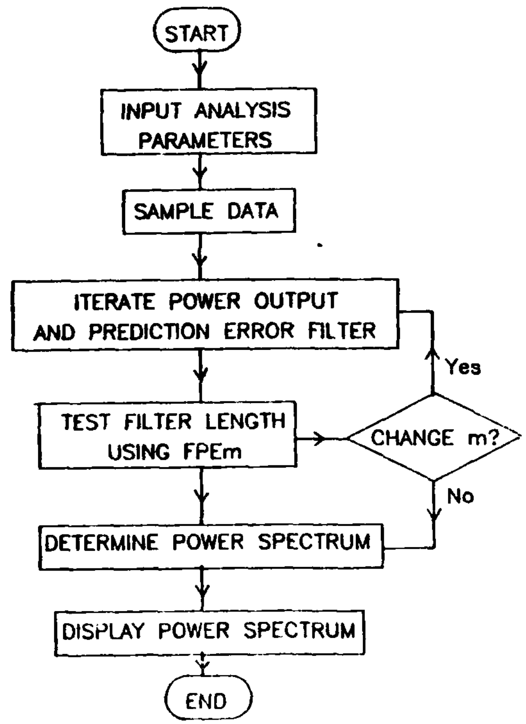


FIGURE 1 STRUCTURE OF MESA PROGRAMS

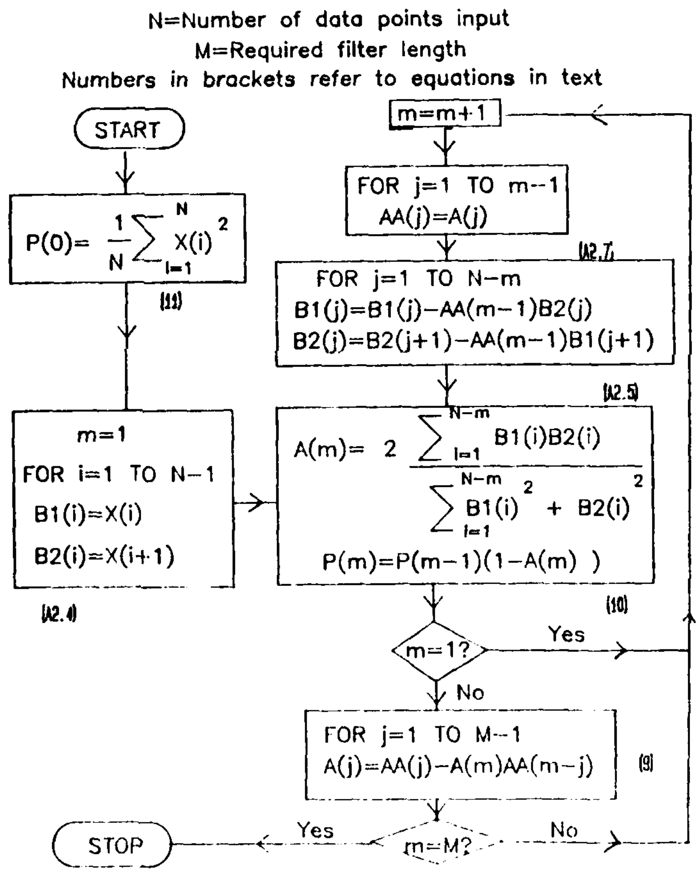


FIGURE 2 FLOW CHART OF ITERATIVE PROCEDURE (REF 5)

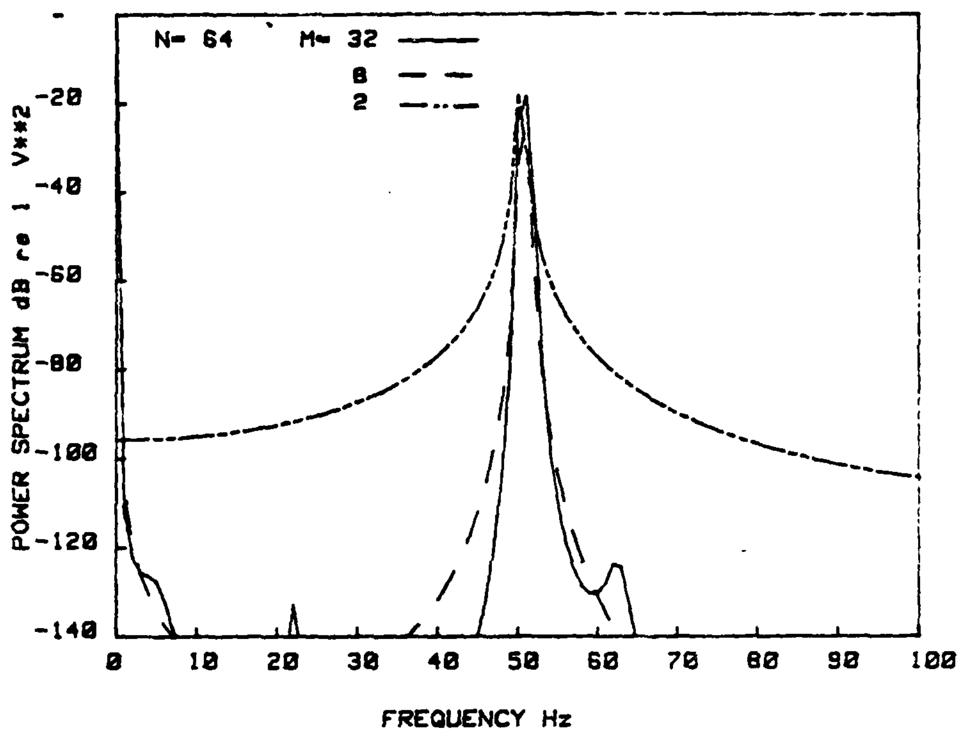


FIGURE 3. VARIATION OF MAXIMUM ENTROPY SPECTRUM WITH NUMBER OF ITERATIONS

(1v rms 50Hz sine wave sampled at 256Hz.)

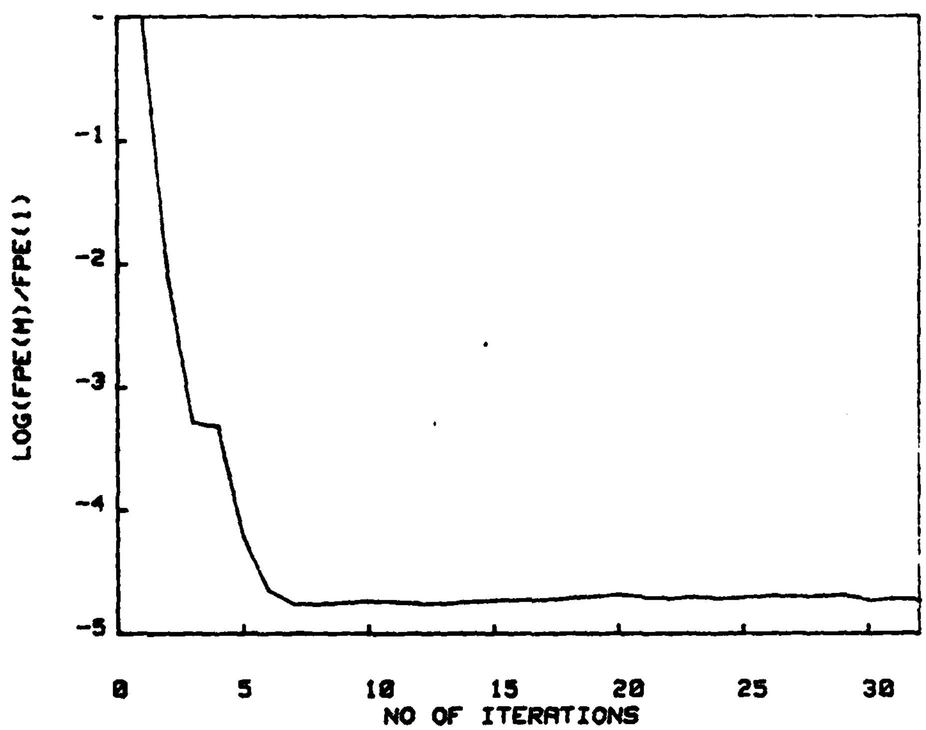


FIGURE 4. VARIATION OF FPE WITH M

(50Hz Sine Wave; N=64; Sampling Frequency 256Hz)

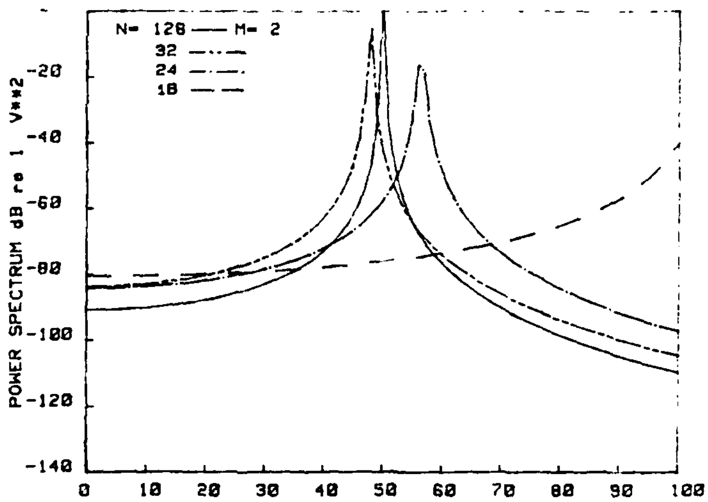


FIGURE 5. VARIATION OF MAXIMUM ENTROPY SPECTRUM
WITH NUMBER OF DATA POINTS
(1v rms 50Hz sine wave sampled at 3200Hz)

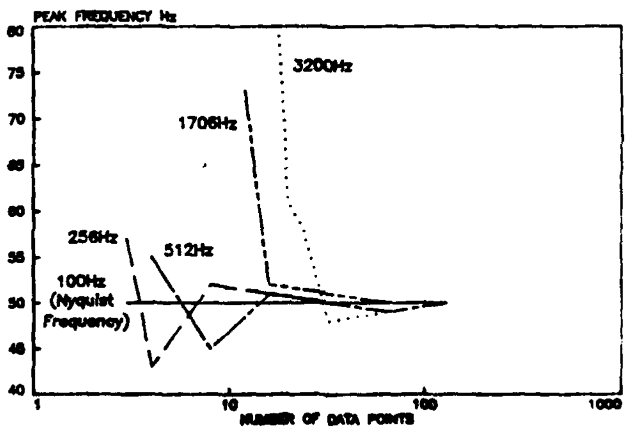


FIGURE 6. VARIATION OF PEAK POSITION WITH NUMBER OF DATA POINTS.

(50Hz Sine Wave; Sampling Frequency as Indicated; M=2)

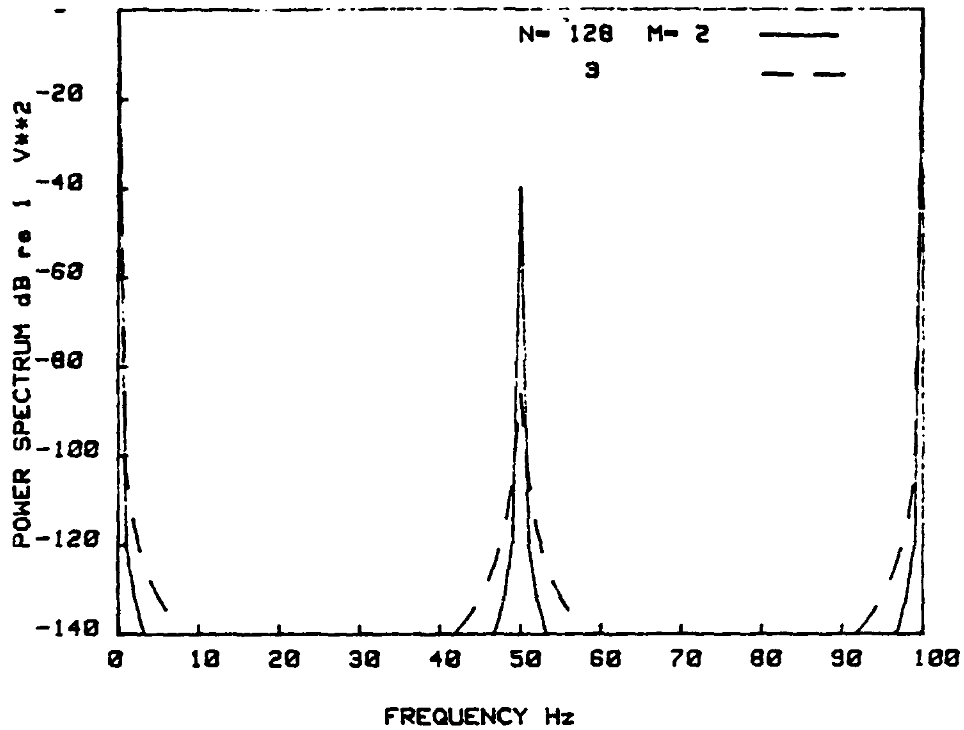


FIGURE 7 MAXIMUM ENTROPY SPECTRUM OF 50Hz SINE WAVE SAMPLED AT NYQUIST FREQUENCY

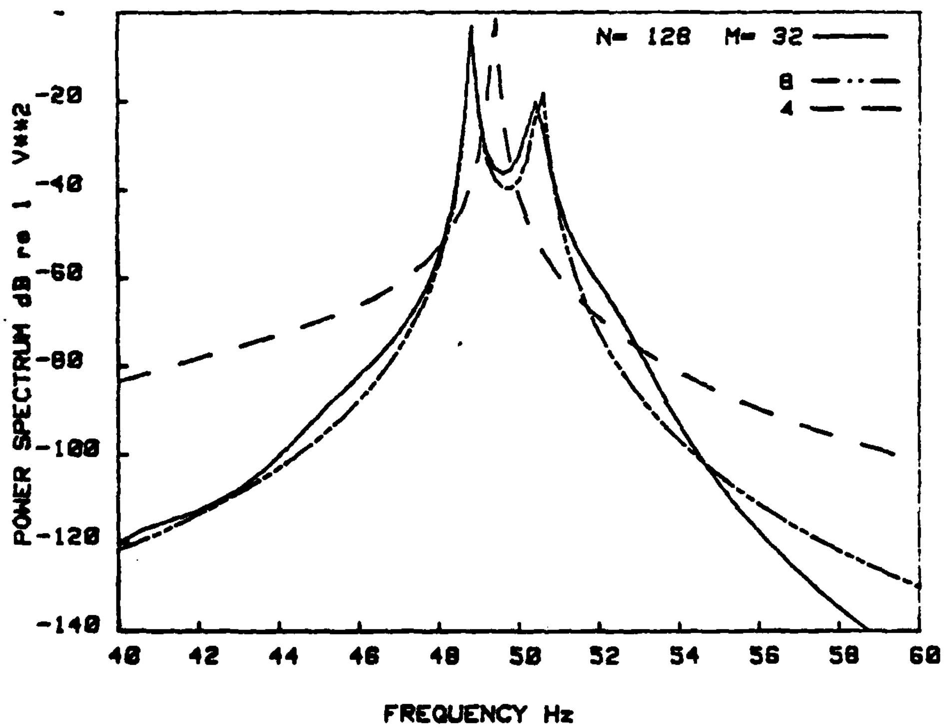


FIGURE 8. MESA PERFORMANCE ON TWO SINE WAVES
1v rms at 49Hz and 50Hz; Sampling Frequency 200Hz

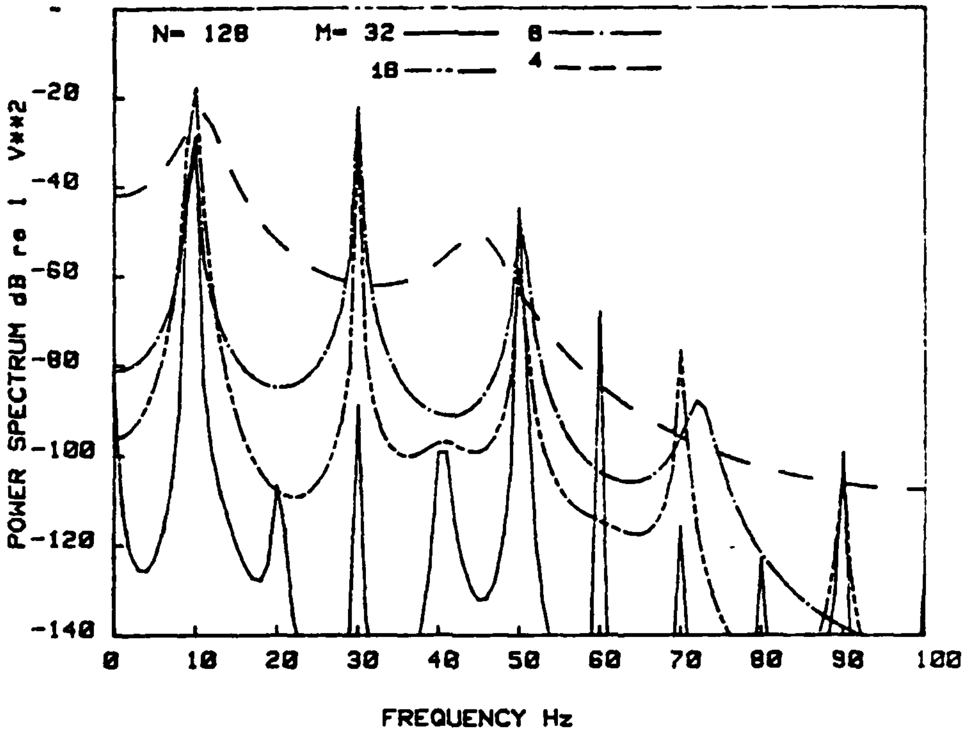


FIGURE 9. MAXIMUM ENTROPY SPECTRUM OF 10Hz SQUARE WAVE

(Sampling Frequency 200Hz)

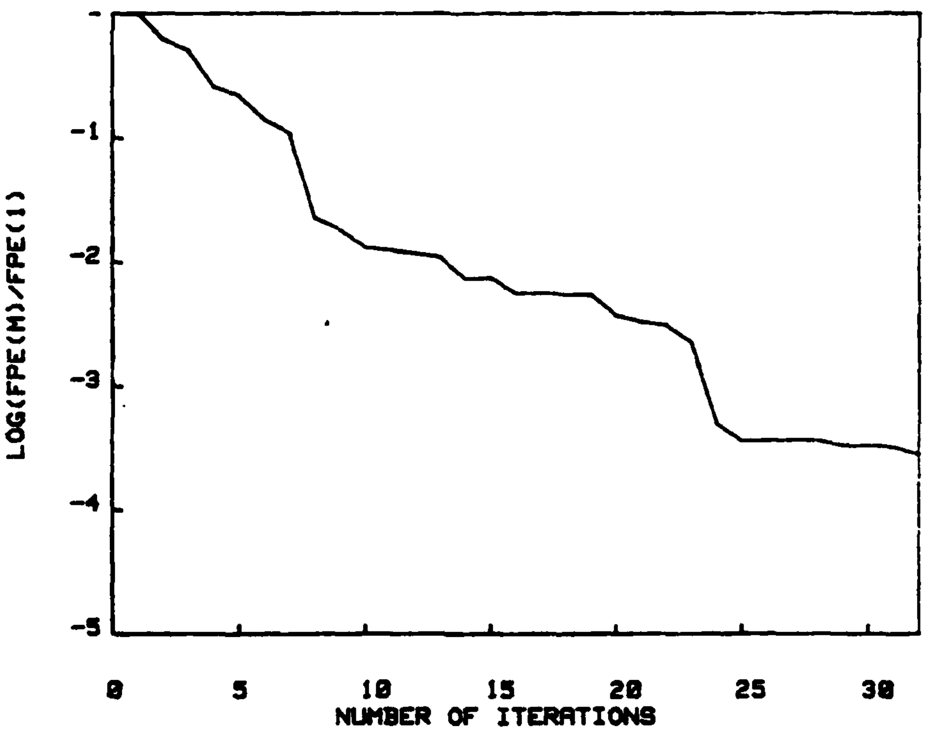


FIGURE 10. VARIATION OF FPE WITH M

(10Hz Square Wave; N=128; Sampling Frequency 200Hz)

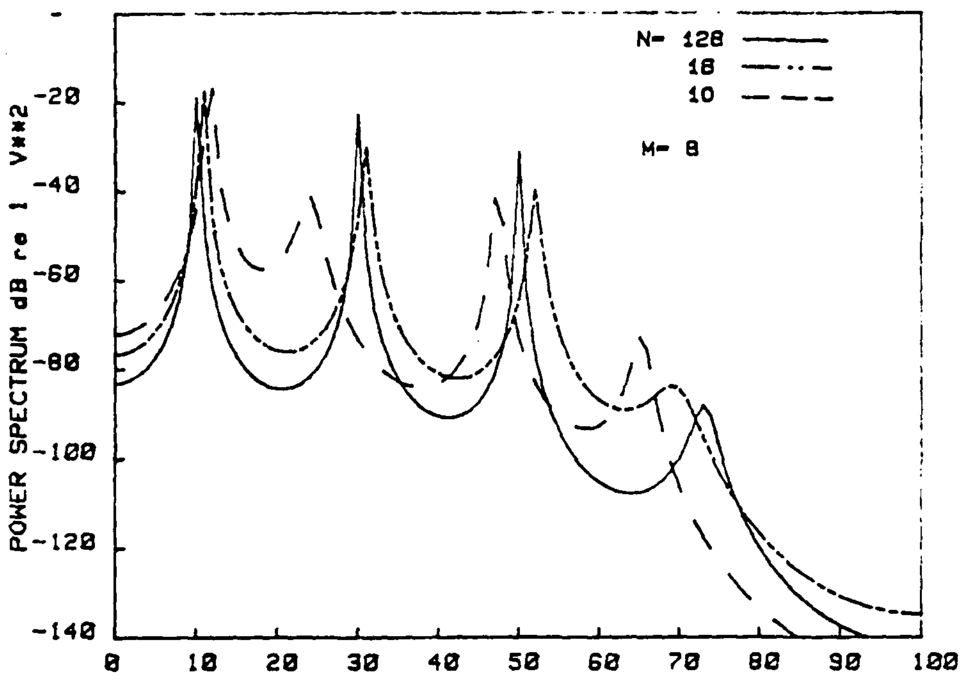


FIGURE 11. VARIATION OF MAXIMUM ENTROPY SPECTRUM
WITH NUMBER OF DATA POINTS
(10Hz Square Wave Sampled at 200Hz)

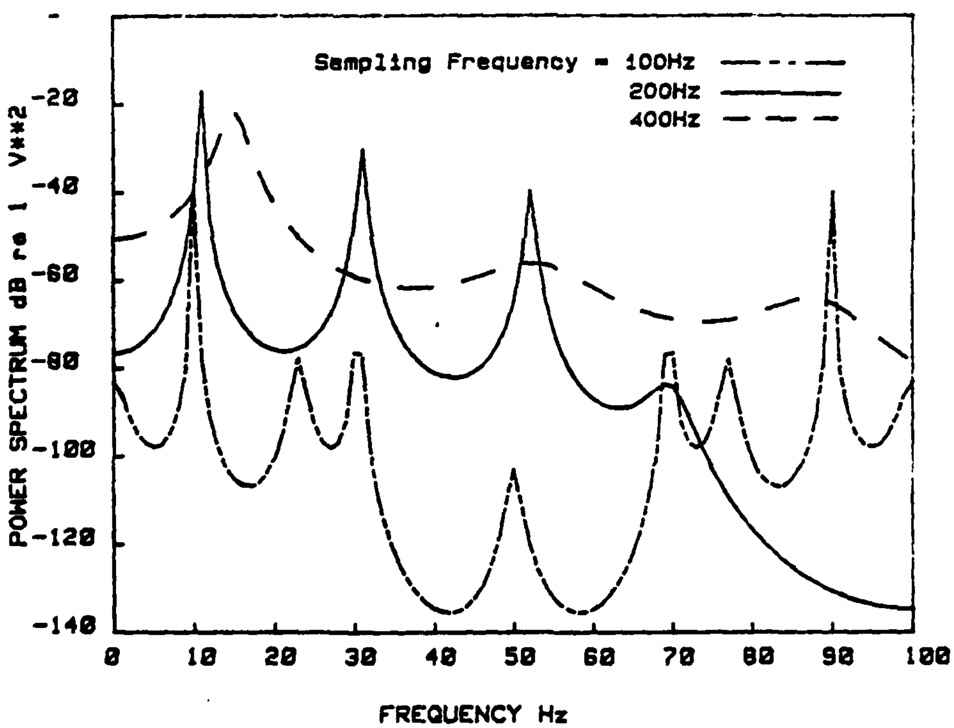


FIGURE 12. VARIATION OF MAXIMUM ENTROPY SPECTRUM
WITH SAMPLING FREQUENCY
(10Hz Square Wave; N=16; M=8)

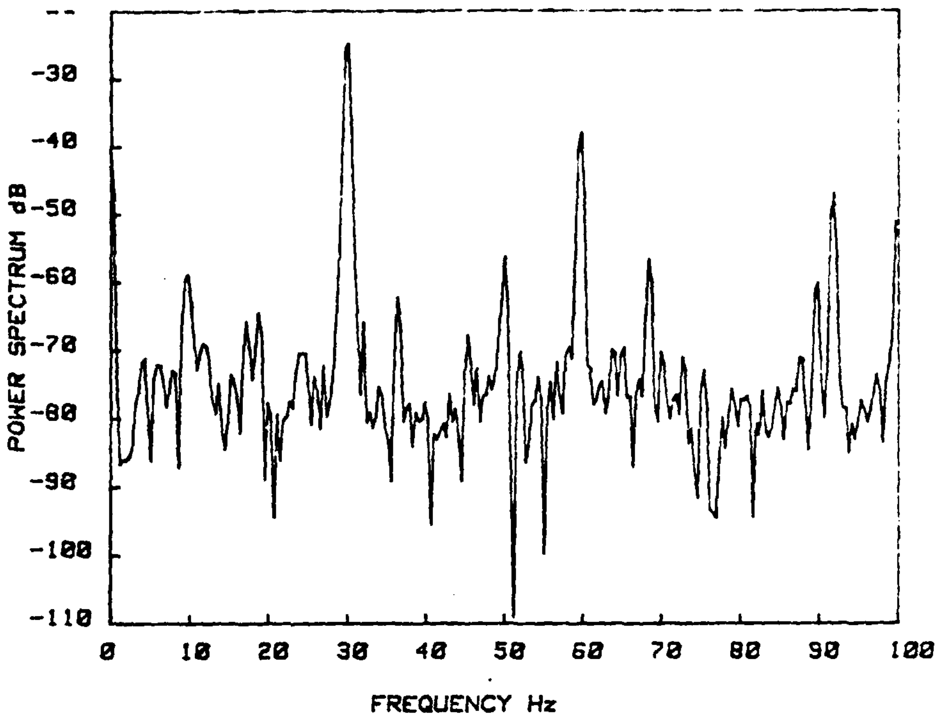


FIGURE 13. EXAMPLE MACHINERY NOISE SPECTRUM
DERIVED USING FFT ON 512 DATA POINTS.

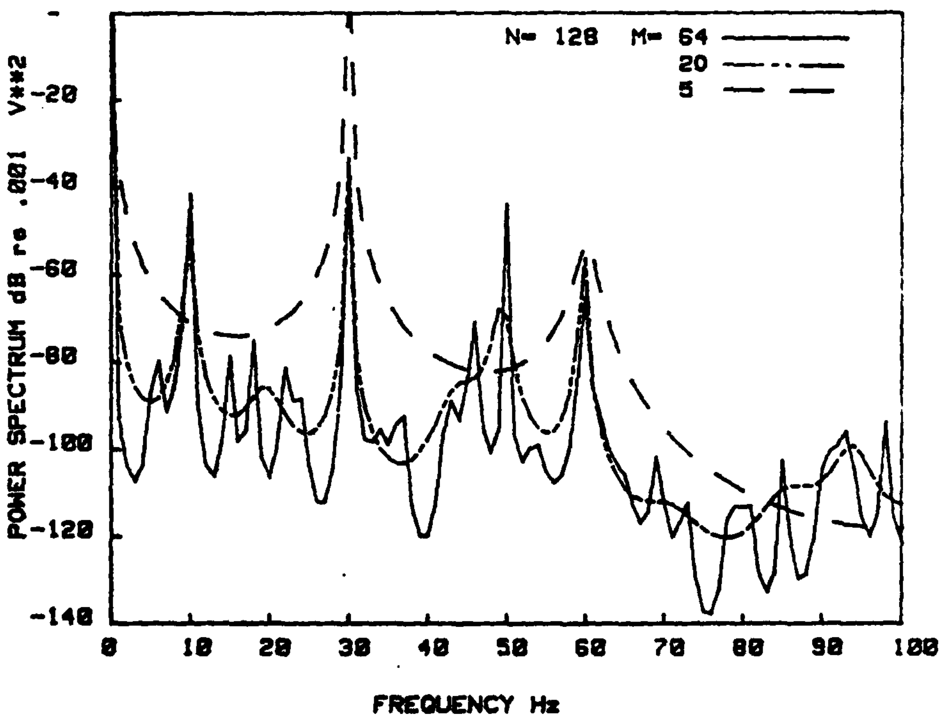


FIGURE 14. MESA PERFORMANCE ON MACHINERY NOISE
Sampling Frequency 200Hz

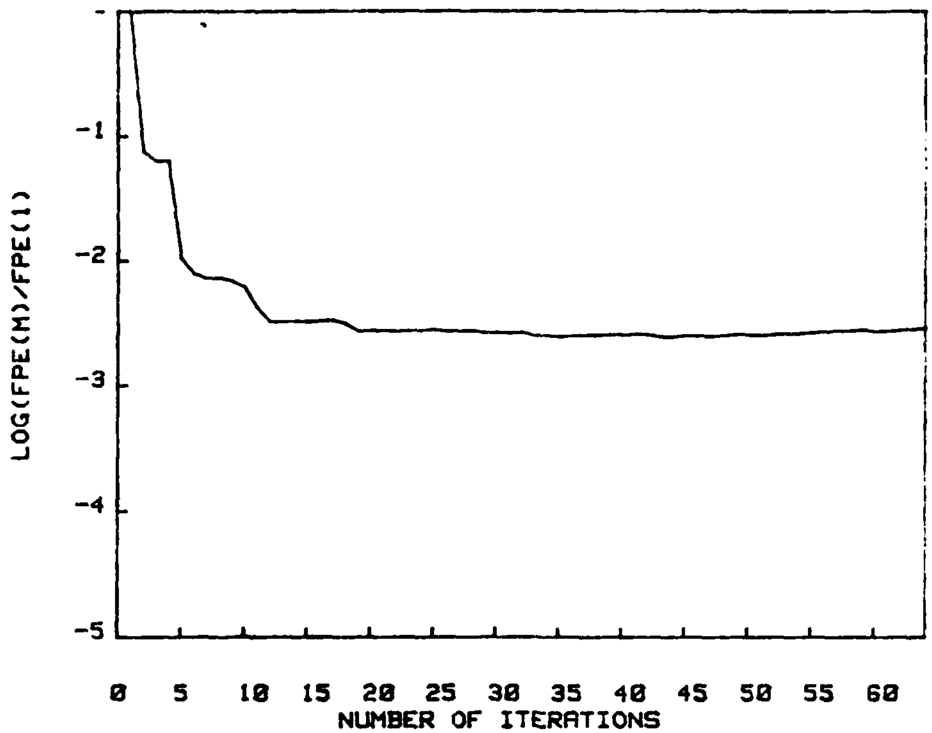


FIGURE 15. FPE CRITERION FOR MACHINERY NOISE EXAMPLE
(N=128; Sampling Frequency 200Hz)

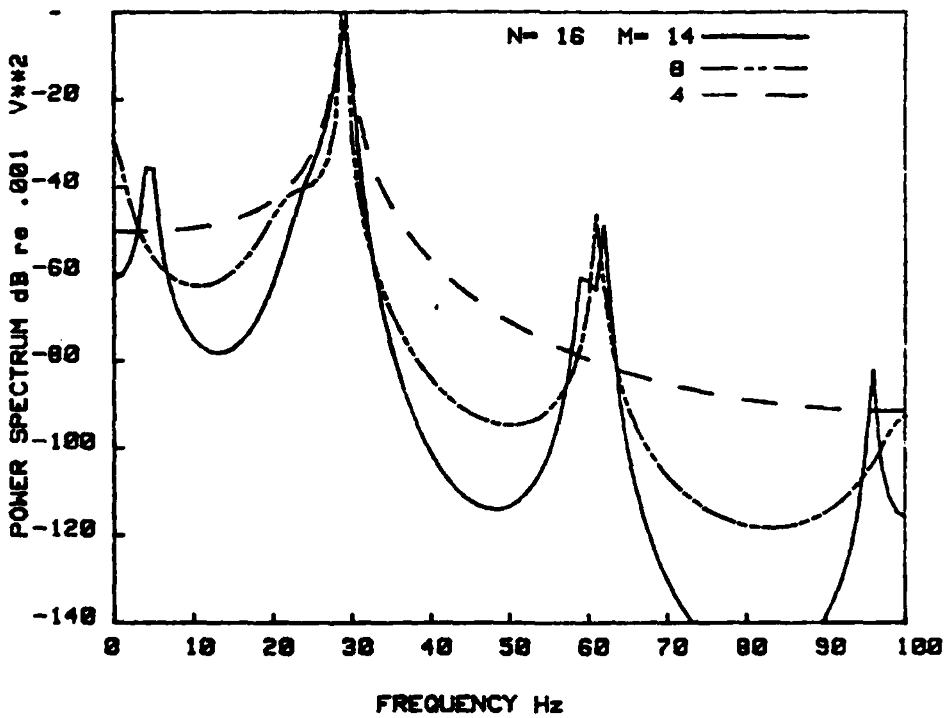
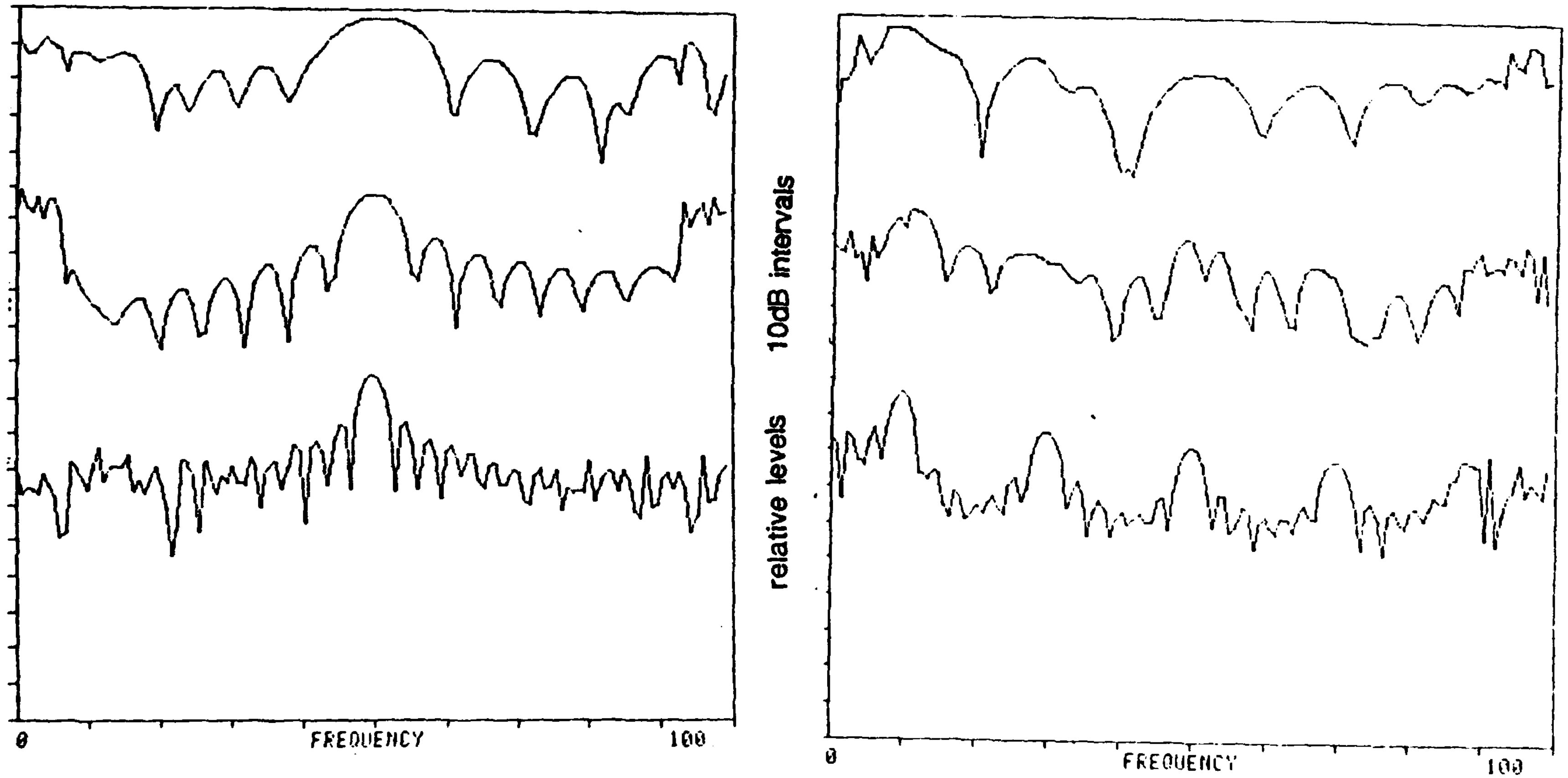


FIGURE 16. MESA PERFORMANCE ON FEW DATA POINTS OF MACHINE NOISE
Sampling Frequency 200Hz



50Hz SINE WAVE

10Hz SQUARE WAVE

200Hz sampling frequency

FIGURE 17 FFT PERFORMANCE ON SMALL DATA SAMPLES

DATE
ILME